

Alinhamento Global de Sequências

Grupo

Daniel Miranda

Valéria Pereira

Como executar

Alinhamento de duas sequências

```
python main.py SEQUENCIA1 SEQUENCIA2
```

ou

```
python main.py -f nome_arquivo_de_input.fasta
```

O arquivo pode ter qualquer extensão, porém seu conteúdo deve estar no formato FASTA.

Essas opções imprimem a pontuação do alinhamento e as sequências alinhadas, caso for desejado visualizar a matriz que resultou no alinhamento, basta adicionar -v na linha de comando

Exemplos:

input:

```
python main.py DRQT DRQTA
```

output:

```
Score: 21
```

```
DRQT-  
DRQTA
```

input:

```
python main.py DRQT DRQTA -v
```

output:

```
      *,  D,  R,  Q,  T,  
* [0 ., 0 ., 0 ., 0 ., 0 .]  
D [0 ., 6 \, 6 _, 6 _, 6 _]  
R [0 ., 6 |, 11 \, 11 _, 11 _]  
Q [0 ., 6 |, 11 |, 16 \, 16 _]  
T [0 ., 6 |, 11 |, 16 |, 21 \]  
A [0 ., 6 |, 11 |, 16 |, 21 |]
```

```
Score: 21
```

```
DRQT-  
DRQTA
```

Alinhamento de múltiplas sequências

A forma de execução é a mesma que a de alinhamento de duas sequências, tendo suporte também para arquivos no formato FASTA, a diferença é o output do script. Como não conseguimos implementar a heurística completa, apenas imprimimos o consenso final do alinhamento, sem remontar as sequências, o modo verboso não imprime nenhuma informação a mais.

Exemplo:

input:

```
python main.py DRQT DRQTA DRQ
```

output:

```
DRQXX
```

Alterando valores do algoritmo

- Para alterar o valor somado em indels, basta abrir o arquivo `.env` e alterar o valor da constante INDEL
- Para alterar o valor de matches e mismatches, basta alterar o dicionário no arquivo `src/blosum.py` os valores contidos nele foram retirados da tabela BLOSUM62